JOURNAL OF THE OPERATIONAL RESEARCH SOCIETY
https://doi.org/10.1080/01605682.2019.1678409

ORIGINAL ARTICLE

Efficiency measurement for hierarchical situations

Wanghong Li[a], Wade D. Cook[a], Zhepeng Li[a] and Joe Zhu[b]

[a]Schulich School of Business, York University, Toronto, Canada; [b]School of Business, Worcester Polytechnic Institute, Worcester, MA, USA

ABSTRACT
The measurement and monitoring of the efficiency of processes in organisations has become an important undertaking in today's competitive environment. A fundamental tool for this undertaking is data envelopment analysis (DEA). The conventional setting for DEA views the decision-making unit (DMU) (school, hospital etc.) as a black box with inputs entering and outputs leaving. The current paper looks at a problem setting somewhat related to a multistage situation but pertaining to a particular form of hierarchical structure. Specifically, we examine a set of electric *power units* that act as sub-units or sub-DMUs, operating under the framework of set of *power plants* that play the role of DMUs. We develop a DEA-like methodology that evaluates, in a two-stage manner, both the efficiencies of the sub-units and of the aggregates of those sub-units (the plants). In so doing, the approach attempts to have the *projected values* of plant-level inputs and outputs match up with the corresponding aggregate values of the sub-unit projections, as is the case prior to projection to the frontier. Since such projections may in fact not *match up* as described, we introduce a goal-DEA methodology to minimise the extent of any failure to achieve this match up.

ARTICLE HISTORY
Received 15 March 2019
Accepted 25 September 2019

KEYWORDS
DEA; efficiency; hierarchy; power plants; sub-units

1. Introduction

The measurement and monitoring of the efficiency of processes in organisations has become an important undertaking in today's competitive environment. Management in those organisations look for ways to improve upon the efficiency of the related processes under their direction. A fundamental tool for this undertaking is data envelopment analysis (DEA), pioneered by Charnes, Cooper, and Rhodes (1978). This methodology is designed to measure the efficiency of a decision-making unit (DMU) in terms of a set of inputs and outputs and *relative* to other similar units. In this sense, DEA aims to identify best practices in the setting under evaluation. Example application areas are schools, hospitals, manufacturing plants, etc.

The conventional setting for DEA views the DMU (school, hospital etc.) as a "black box" with inputs entering and outputs leaving. Inputs might take the form of staff, budgets, or environmental factors; outputs can be products, numbers of customers served, and so on. The ideas behind DEA have been expanded in many directions, going well beyond the initial constant returns to scale (CRS) model of Charnes et al. (1978). These developments include the methodological extension to other model forms, and specifically the variable returns to scale (VRS) model by Banker, Charnes, and Cooper (1984), the additive model of Charnes, Cooper, Golany, Seiford, and Stutz (1985), and the slacks-based model (SBM) by Tone (2001). One important direction where there is a large body of the literature is that involving multistage processes such as one encounters in supply chains. See Liang, Cook, and Zhu (2010) for a survey of some of that literature. In a recent paper by Li, Cook, Li, and Liang (2017), for example, the decision-making units are various regions or provinces in China that are looked at from a two-stage perspective. The first stage represents the economic development of regions, while the second stage describes the recycling of hazardous waste resulting from the economic development. A number of surveys on DEA have been carried out including Cook and Seiford (2009).

The current paper looks at a problem setting somewhat related to multistage situations but pertaining to a particular form of hierarchical structure. Specifically, we examine 18 electric *power units* that act as sub-units or sub-DMUs, operating under the framework of 9 *power plants* that play the role of DMUs. We approach efficiency evaluation at the plant level by viewing it as the sum of the power units that make up the plant. This idea initially appeared in two earlier papers by Chai, Cook, Green, and Doyle (1998) and Cook and Green (2005). We point out that we use a subset of the

CONTACT Wade D. Cook wcook@schulich.yorku.ca Schulich School of Business, York University, Toronto, M3P 4C3 Canada

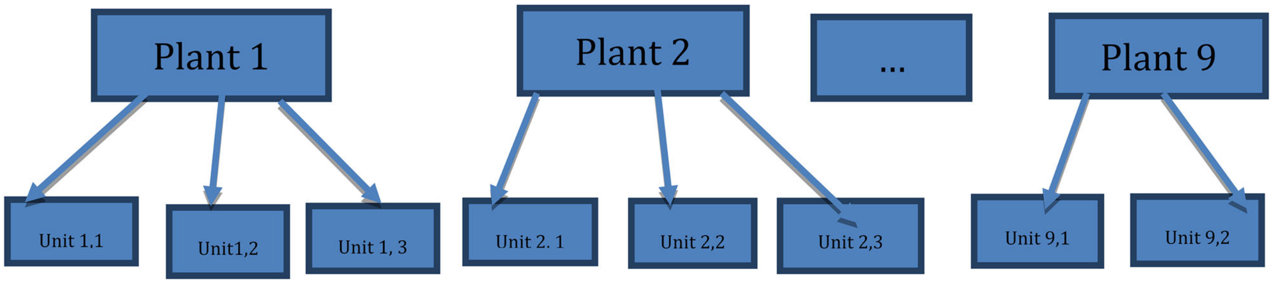

Figure 1. A hierarchical structure in DEA.

original set of 10 plants and 42 sub-units described in the Chai et al. (1998) and Cook and Green (2005) papers. It is observed that another important and related problem is that involving parallel process situations, particularly that discussed in Kao (2009), and in somewhat similar vein, the earlier work of Castelli, Pesenti, and Ukovich (2004) proposed a model to measure the performance of a hierarchical system with one and two-level where the units at the higher level can have common subordinate units at the lower level. The models of Castelli are designed to measure efficiency only at the level of the whole DMU.

Kao (2015) explicitly examines hierarchical systems in which each DMU has the same number of first level sub-units performing distinctive functions, although he shows that this is not actually a strict requirement. It is assumed that the outputs of any intermediate sub-unit come completely from its subordinate sub-units, meaning that the intermediate unit does not produce outputs itself. The model is designed to optimise the efficiency (the ratio of weighed aggregate outputs to weighted aggregate inputs) of the DMU (top) level in the hierarchy, subject to the constraints at the sub-unit levels (using the optimal DMU-level output and input weights). These latter sub-unit output to input ratios are taken as the sub-unit efficiencies. Thus, we can refer to the subordinate or sub-unit efficiency scores as being “derived” from the optimised DMU-level scores. In so doing the same multipliers are applied at the level of any sub-unit as at the top (DMU) level. This is as discussed in the relational model of Kao (2009). It is pertinent to point out that model (3.4) in Cook and Green (2005) treats efficiencies of sub-units as being derived from the optimised top level efficiency model in the same manner as in Kao (2015).

In the current paper, we revisit the power plant problem of Cook and Green (2005) involving the efficiency measurement of power plants. Section 2 describes the problem of efficiency evaluation in a set of power units. Section 3 presents a hierarchical DEA models for measuring efficiency at two levels, and the application of this model to the power plant problem is presented in Section 4. It is important to point out that the primary difference between the model herein versus that of the earlier works of Kao (2015), Castelli et al. (2004) and others is that we optimise the efficiencies of sub-units as well as at the top DMU level. This approach has benefits and disbenefits to be discussed below. Conclusions and further research direction are presented in Section 5.

2. Hierarchical efficiency involving power plants

In this paper, we examine data on the performance of a set of thermal power units of varying ages, capacities and fuel types as described in Cook and Green (2005). For completeness, we paraphrase a portion of the discussion in that earlier paper.

Historically, the efficiency of a plant or sub-unit has been represented by the ratio of total production in megawatt hours to total annual expenditure. The earlier papers by Cook and Green (2005) and Chai et al. (1998) expanded the conventional definition of productivity to include other factors. Management is interested not only in the efficiency of the power plants but as well that of each sub-unit within any plant. This two-level evaluation provides for a description of the extent of improvement necessary at both the plant and sub-unit levels (see Figure 1).

As discussed in the previous paper of Cook and Green (2005), management’s skill is a factor that should come into consideration in evaluating power plant efficiency. It can be argued that a tangible form of management skill is the quality of maintenance carried out in a power unit. While maintenance cost clearly reflects the level of work done, the tangible consequences of this activity appear in the form of power outages and equipment failure deratings. We explain these two concepts.

An *outage* is a situation in which a unit is shut down, hence it is not generating electric power. Types of outages include:

- Planned outage, which is scheduled downtime.
- Maintenance outage, a form of scheduled down time, but for more minor, i.e., routine maintenance.

- Forced outage, which is unscheduled and generally caused by equipment failure, environmental requirements, or other unforeseen incidents. There is generally some prior warning for this type of shutdown, and some delay is possible.
- Sudden outage, which is a forced outage with no prior warning.

While it can be argued that operating hours essentially captures all forms of outages, it must be recognised that there is a difference between taking a unit out of service on a scheduled basis at non-peak times, versus sudden brownouts or blackouts. The latter ignite public opinion, interrupt business operations, and generally reflect negatively on the organisation. Thus, such outages should play a direct role in any measure of efficiency.

A *derating* is a reduction in power unit capacity, wherein the unit operates at only a fraction of its available full capacity. As with outages, the prime example of this is equipment failure. In other cases, derating can occur because of there being excess environment pollution, generally in the form of SO_2 emissions.

The previous study by Chai et al. (1998) used a two-stage process to measure performance, first at the level of the power unit and second at the plant level. This was then followed by adjustments to sub-unit (level 1) ratings, utilising the level 2 rating of the respective plants. Cook and Green (2005) argued that there were two disadvantages of that earlier two stage approach. First, the earlier paper modelled damaging gas emissions by way of the *frequency* of environmental deratings per year, as opposed to some function of the level of the SO_2, above or below a threshold. Second, since the environmental variable only applies at the plant level, the earlier model of Cook et al. (1998) chose to assume that power units should be equally penalised. Hence, there was an issue as to how to appropriately capture environmental damage at level 1.

Cook and Green (2005) presented a different version of the DEA model that viewed both levels in the hierarchy simultaneously, by focusing the optimisation at the plant or DMU level while insuring *feasibility* at the power unit level. It is important to point out that their model does not necessarily render the sub-unit efficiency scores "optimal," but rather they are simply feasible. An outstanding issue then remains as to how to identify some form of optimal measure of efficiency at the two hierarchy levels simultaneously. In the section to follow, we present an augmented type of the DEA model that attempts to evaluate each plant as well as the sub-units within those plants.

3. Modelling efficiency in a hierarchy

Before proceeding, it is useful to present the multiplier and envelopment forms of the standard CRS models of Charnes et al. (1978), namely (3.1) and (3.2), respectively. This serves to set the stage for the two-level model discussed below. We point out that (3.1) is the linear transformation of the usual fractional programming version of the Charnes et al. (1978) model:

$$\begin{aligned}
&\text{Max} \sum_{r=1}^{R} \mu_r y_{rj_o} \\
&\text{Subject to :} \\
&\quad \sum_{i=1}^{I} \upsilon_i x_{ij_o} = 1 \\
&\quad \sum_{r=1}^{R} \mu_r y_{rj} - \sum_{i=1}^{I} \upsilon_{ij} x_{ij} \le 0, \forall j \\
&\quad \mu_r, \upsilon_i \ge \varepsilon, \forall i, r
\end{aligned} \tag{3.1}$$

$$\begin{aligned}
&\text{Min} \left[\theta_{j_o} - \varepsilon \left(\sum_{i \in I} s_{ij_o} + \sum_{r \in R} s_{rj_o} \right) \right] \\
&\text{Subject to :} \\
&\quad -\sum_{j=1}^{n} \lambda_j x_{ij} + \theta_{j_o} x_{ij_o} - s_{ij_o} = 0, \forall i \\
&\quad \sum_{j=1}^{n} \lambda_j y_{rj} - s_{rj_o} = y_{rj_o}, \forall r \\
&\quad \lambda_j \ge 0
\end{aligned} \tag{3.2}$$

Note that we have introduced the infinitesimal ε in model (3.1), meaning that slacks are explicitly accounted for in (3.2).

The methodology presented in the current paper revisits the problem of efficiency measurement at two levels in a hierarchy, pertaining to electric power generation; we again point out that level 1 relates to the power units, while level 2 concerns the power plants in which the units are located. In modelling efficiency in this area, we wish to adhere to two basic principles, namely:

Principle 1: The inputs X_{ik_o} and outputs Y_{rk_o} for a given level 2 DMU (power plant) k_o are obtained by aggregating, across all power units within that level 2 DMU, the corresponding level1 inputs and outputs, i.e., $X_{ik_o} = \sum_{jk_o \in J_{k_o}} x_{ijk_o}$ and $Y_{rk_o} = \sum_{jk_o \in J_{k_o}} y_{rjk_o}$.

Principle 2: In evaluating the input-oriented efficiency of any level 2 DMU (power plant) k_o, input reduction is done in such a way that the post-projection input and output values at level 2 are given by the aggregates of the corresponding post-projection level 1 input and output values.

Regarding Principle 1, it is possible that there may be inputs and outputs that are generated at level 2 (the power plant), in addition to those appearing at level 1. In this case, we assume that such level 2 factors can be split up and allocated to the level 1 sub-DMUs that lie within that level 2 unit.

Regarding Principle 2, it must be reiterated that the level 2 DMU is *explicitly defined* as being the sum of the sub-DMUs within that DMU. Specifically, as given above, each input X_{ik} and output Y_r held by a level 2 DMU is the sum of the corresponding values of that variable held by the sub-DMUs within that DMU. Hence, when the inputs of a DMU are reduced, while being projected to the level 2 frontier, the resulting *projected DMU* must also be defined as the sum of the *projected sub-DMUs*. It is important to recognise that Principle 2 was not imposed in the previous work presented in Cook et al. (1998) and Cook and Green (2005).

To facilitate the development of our model, we utilise the following notation, for modelling efficiency at the two levels in the hierarchy.

K is the number of DMUs in level 2 of the hierarchy (Power Plants). n is the number of DMUs in level 1 of the hierarchy (Power Units).

k is the power plant index.

I is the number of inputs.

R is the number of outputs.

J_k is the group of sub-units (power units) in DMU k.

j_k is the index for sub-units in group J_k.

$x^k_{ij_k}$ is the quantity of input i held by sub-unit j_k in power plant k.

$y^k_{rj_k}$ is the quantity of output r held by sub-unit j_k in power plant k.

θ_{k_o} is the variable representing the efficiency measure of DMU (power plant) k_o.

$\lambda_j(j_{k_o})$ is the variable used in the envelopment model (3.5) below, representing the weight attached to sub-unit j_k, and identifying which sub-units are on the frontier.

$s_{ij_{k_o}}$ is the slack variable for input i in sub-unit j_{k_o}, and representing the distance from the (weak) projected point $\theta_{k_o}x_{ij_{k_o}}$ to the frontier.

S_{ik_o} is the slack variable for input i in DMU (power plant) k_o, and representing the distance from the (weak) projected point $\theta_{k_o}X_{ik_o}$ to the frontier.

$s_{rj_{k_o}}, S_{rk_o}$ are slacks for the output dimension r.

μ_r is the variable used in the multiplier model (3.3) below, representing the weight attached to output r in sub-unit j_{k_o}

υ_i is the variable used in the multiplier model (3.4) below, representing the weight attached to input i in sub-unit j_{k_o}

$d_i^-, d_i^+, d_r^-, d_r^+$ are goal achievement variables for model (3.5′).

To model the two-level problem where we wish to activate the above principle #2, the conventional model (3.2) does not capture the idea of reducing inputs for all members j_{k_o} of a set J_{k_o}, simultaneously, and will often encounter infeasibilities. To illustrate, consider Figure 2.

In this example, let us assume that F and E are subunits and are members of the same power plant. If we wish to treat these two units as projecting toward the frontier, *simultaneously*, and allowing for the fact that they may possibly receive different minimum θ-values, then we would end up declaring projected points E′ and F′ as the frontier positions that the two sub-DMUs should aim for, in order to be deemed efficient. Note that for E to project to E′, λ_A and λ_B will be positive and all other λ_j will be 0. However, in projecting F to F′, λ_C and λ_B must be positive, and all other λ_j will be set to 0. Clearly, the requirements of the λ_j for points F′ and E′ are inconsistent, meaning that if simultaneous projection of the two sub-DMUs F and E is to occur, the conventional model (3.2) is *infeasible*.

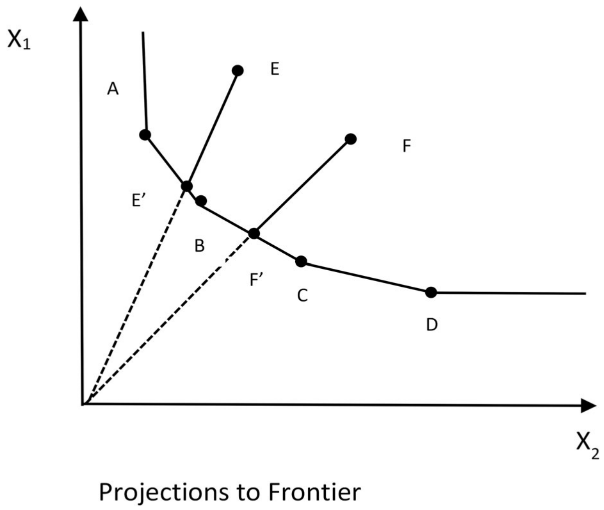

Figure 2. Projection to the frontier.

If, however, F and E need to be projected toward the frontier simultaneously, then it would appear that each of these sub-DMUs should have its *own* set of λ_j, i.e., $\lambda_j(F)$ and $\lambda_j(E)$, or more generally sets $\lambda_j(j_{k_o})$ for each sub-unit j_{k_o} in a set J_{k_o}. We propose replacing the conventional model (3.2) by the following two-stage extension of that model.

Stage 1: In this stage, we derive input-oriented efficiencies of the level 2 DMUs, utilising aggregates of the inputs and outputs as defined under principal 1 above. Viewed from a multiplier perspective the efficiency scores for DMU k_o would be derived utilising model (3.3)

$$\begin{aligned}
&\text{Max} \sum_{r=1}^{R} \mu_r Y_{rk_o} \\
&\text{Subject to :} \\
&\quad \sum_{i=1}^{I} \upsilon_i X_{ik_o} = 1 \\
&\quad \sum_{r=1}^{R} \mu_r Y_{rk} - \sum_{i=1}^{I} \upsilon_i X_{ik} \leq 0, \forall k, \\
&\quad \sum_{r=1}^{R} \mu_r y_{rj} - \sum_{i=1}^{I} \upsilon_i x_{ij} \leq 0, \forall j \\
&\quad \mu_r, \upsilon_i \geq \varepsilon, \forall i, r
\end{aligned} \tag{3.3}$$

The dual of (3.3) is given by (3.4).

$$\text{Min}\left[\theta_{k_o} - \varepsilon\left(\sum_{i\in I} S_{ik_o} + \sum_{r\in R} S_{rk_o}\right)\right]$$

Subject to :

$$-\sum_{j=1}^{n}\lambda_j^1 x_{ij} - \sum_{k=1}^{K}\lambda_k^2 X_{ik} + \theta_{k_o}X_{ik_o} - S_{ik_o} = 0, \forall i$$

$$\sum_{j=1}^{n}\lambda_j^1 y_{rj} + \sum_{k=1}^{K}\lambda_k^2 Y_{rk} - S_{rk_o} = Y_{rk_o}, \forall r$$

$$\lambda_j^1, \lambda_k^2 \geq 0, \forall j, k, S_{rk_o}, S_{ik_o} \geq 0, \forall r, i \quad (3.4)$$

It is model (3.4) that we implement in Section 4 for each DMU (plant) k_o.

Stage 2: Consider now model (3.5).

$$\text{Min}\sum_{j_{k_o}\in J_{k_o}}\left[\theta_{j_{k_o}} - \varepsilon\left(\sum_{i\in I} s_{ij_{k_o}} + \sum_{r\in R} s_{rj_{k_o}}\right)\right] \quad (3.5a)$$

Subject to:

$$-\sum_{j=1}^{n}\lambda_j(j_{k_o})x_{ij} + \theta_{j_{k_o}}x_{ij_{k_o}} - s_{ij_{k_o}} = 0, \forall i \in I, \forall j_{k_o} \in J_{k_o} \quad (3.5b)$$

$$\sum_{j=1}^{n}\lambda_j(j_{k_o})y_{rj} - s_{rj_{k_o}} = y_{rj_{k_o}}, \forall r \in R, \forall j_{k_o} \in J_{k_o} \quad (3.5c)$$

$$\sum_{j_{k_o}\in J_{k_o}}(\theta_{j_{k_o}}x_{ij_{k_o}} - s_{ij_{k_o}}) = \hat{\theta}_{k_o}X_{ik_o} - \hat{S}_{ik_o}, \forall i \in I \quad (3.5d)$$

$$\sum_{j_{k_o}\in J_{k_o}}(y_{rj_{k_o}} + s_{rj_{k_o}}) = Y_{rk_o} + \hat{S}_{rk_o}, \forall r \in R \quad (3.5e)$$

$$\lambda_j(j_{k_o}) \geq 0, \forall j, j_{k_o} \in J_{k_o}; s_{ij_{k_o}}, s_{rj_{k_o}} \geq 0, \forall i, r, j_{k_o}; \theta_{k_o} \text{ unrestricted in sign} \quad (3.5f)$$

This model derives an optimal efficiency score $\hat{\theta}_{j_{k_o}}$ for each sub-unit $j_{k_o} \in J_{k_o}$. It is important to emphasise here that model (3.5) is an attempt to allow for target setting for both sub-units (power units) and power plants. This is in contrast to Kao (2015) where generally targets are set only for aggregates or the DMU level. The model is intended to exhibit two important properties. *First,* as discussed above, it is desirable that the level 2 DMU k_o be defined (in terms of quantities of inputs and outputs held) as the sum of the level 1 sub-DMUs j_{k_o} within that DMU, *post-projection* to the efficient frontier as is the case *pre-projection*. Constraints (3.5d, 3.5e) are meant to meet this requirement. The right sides represent the projected values of the inputs and outputs for (level 2) DMU k_o to the frontier, while the left-hand side represents the sum of projected inputs and outputs to the *set* of frontiers for the sub-DMUs $j_{k_o} \in J_{k_o}$. (We point out below, however, that such a requirement may fail to be met).

The *second* property is the presence of a separate set of vectors $\{\lambda_j(j_{k_o})\}$ for the members j_{k_o} in J_{k_o}, the collection of sub-DMUs in plant k_o. As described earlier, this permits the set $\{\hat{\theta}_{j_{k_o}}\}$ of sub-DMU efficiency scores to be derived *simultaneously* rather than independently of one another. This approach provides a mechanism for imposing conditions (3.5d) in one problem setting that independent projections of sub-DMUs would not permit.

We note that the objective function $\sum_{j_{k_o}\in J_{k_o}}[\theta_{j_{k_o}} - \varepsilon(\sum_{i\in I} s_{ij_{k_o}} + \sum_{r\in R} s_{rj_{k_o}})]$ in (3.5) is intended to not only identify the (weak) efficiency scores $\hat{\theta}_{j_{k_o}}$ for each of the sub-DMU $sj_{k_o} \in J_{k_o}$, but as well any positive slacks $\hat{s}_{ij_{k_o}}, \hat{s}_{rj_{k_o}}$ representing the additional reductions in inputs and increases in outputs that are required to bring those sub-DMUs to the frontier.

We prove the following theorem.

Theorem 3.1: A DMU (power plant) k_o exhibits *strong* efficiency if and only if each of the sub-DMUs $j_{k_o} \in J_{k_o}$ also exhibits strong efficiency.

Proof:

1. If all sub-units $j_{k_o} \in J_{k_o}$ of a given DMU k_o are strongly efficient, then regarding (3.5d), all sub-unit scores $\hat{\theta}_{j_{k_o}} = 1$, and all slacks $s_{ij_{k_o}}, s_{rj_{k_o}} = 0$, meaning that the left-hand side of (3.5) becomes $\sum_{j_{k_o}\in J_{k_o}}(\theta_{j_{k_o}}x_{ij_{k_o}} - s_{ij_{k_o}}) = \sum_{j_{k_o}\in J_{k_o}} x_{ij_{k_o}}$, hence $\sum_{j_{k_o}\in J_{k_o}} x_{ij_{k_o}} = X_{ik_o}$. But, if the right-hand side expression $\hat{\theta}_{k_o}X_{ik_o} - \hat{S}_{ik_o}$ becomes X_{ik_o} then it must be that $\hat{\theta}_{ik_o} = 1$ and the slacks $\hat{S}_{ik_o} = 0$. Thus, on the input side DMU k_o is strongly efficient. On the output side (constraint (3.5e)), the left-hand side becomes $\sum_{j_{k_o}\in J_{k_o}} y_{rj_{k_o}}$ which is Y_{rk_o}, meaning that all $S_{rk_o} = 0$, hence on the output side DMU k_o is strongly efficient.
2. If at least one sub-unit $j_{k_o} \in J_{k_o}$ is not strongly efficient, then on the input side (constraint (3.5d)), either $\hat{\theta}_{j_{k_o}} < 1$ or at least one slack $s_{ij_{k_o}} > 0$, meaning that the left side of (3.5d) is *strictly less than* X_{ik_o}. Thus, the right side of (3.5d) must be strictly less than X_{ik_o}, which can only be true if the DMU in question is inefficient. On the output side an argument similar to that given above leads to the same result.

This proves the theorem.

3.1. A problem of infeasibility

Model (3.5) above can be infeasible in certain circumstances, as will be demonstrated in the following section. Specifically, the level 1 efficiencies provided by the left sides of constraints (3.5d) and (3.5e) can fail to *match up* with the right sides that represent the level 2 scores and slacks. In this respect, the expressions in (3.5d) and (3.5e) should be treated as targets or *goals* we seek to meet, rather than constraints that *must* be met. Let us introduce two sets of decision variables or goal achievement variables $\{d_i^-, d_r^-\}, \{d_i^+, d_r^+\}$, to represent the under achievement and over achievement, respectively, of the goals. Now, replace constraints (3.5d) and (3.5e) by expressions:

$$\sum_{j_{k_o}\in J_{k_o}} (\theta_{j_{k_o}} x_{ij_{k_o}} - s_{ij_{k_o}}) - \hat{\theta}_{k_o} X_{ik_o} + \hat{S}_{ik_o} + d_i^- - d_i^+ = 0,\ \forall i \in I \quad (3.5'd)$$

$$\sum_{j_{k_o}\in J_{k_o}} (y_{rj_{k_o}} + s_{rj_{k_o}}) - Y_{rk_o} - \hat{S}_{rk_o} + d_r^- - d_r^+ = 0,\ \forall r \in R \quad (3.5'e)$$

Furthermore, we replace the objective function (3.5a) by

$$\text{Min} \sum_{r\in R} \frac{d_r^- + d_r^+}{\bar{Y}_{rk}} + \sum_{i\in I} \frac{d_i^- + d_i^+}{\bar{X}_{ik}} + \sum_{j_{k_o}\in J_{k_o}} \left[\theta_{j_{k_o}} - \varepsilon\left(\sum_{i\in I} s_{ij_{k_o}} + \sum_{r\in R} s_{rj_{k_o}}\right)\right] \quad (3.5'a)$$

Note that we recommend scaling the achievement variables by the average of the level 2 inputs and outputs across all level 2 DMUs (plants) $k \in K$. Thus, we recommend replacing model (3.5) by the goal achievement model (3.5′)

$$\text{Min} \sum_{r\in R} \frac{d_r^- + d_r^+}{\bar{Y}_{rk}} + \sum_{i\in I} \frac{d_i^- + d_i^+}{\bar{X}_{ik}} + \sum_{j_{k_o}\in J_{k_o}} \left[\theta_{j_{k_o}} - \varepsilon\left(\sum_{i\in I} s_{ij_{k_o}} + \sum_{r\in R} s_{rj_{k_o}}\right)\right] \quad (3.5'a)$$

Subject to:

$$-\sum_{j=1}^{n} \lambda_j(j_{k_o}) x_{ij} + \theta_{j_{k_o}} x_{ij_{k_o}} - s_{ij_{k_o}} = 0, \forall i \in I, \forall j_{k_o} \in J_{k_o} \quad (3.5'b)$$

$$\sum_{j=1}^{n} \lambda_j(j_{k_o}) y_{rj} - s_{rj_{k_o}} = y_{rj_{k_o}}, \forall r \in R, \forall j_{k_o} \in J_{k_o} \quad (3.5'c)$$

$$\sum_{j_{k_o}\in J_{k_o}} (\theta_{j_{k_o}} x_{ij_{k_o}} - s_{ij_{k_o}}) - \hat{\theta}_{k_o} X_{ik_o} + \hat{S}_{ik_o} + d_i^- - d_i^+ = 0, \forall i \in I \quad (3.5'd)$$

$$\sum_{j_{k_o}\in J_{k_o}} (y_{rj_{k_o}} + s_{rj_{k_o}}) - Y_{rk_o} - \hat{S}_{rk_o} + d_r^- - d_r^+ = 0, \forall r \in R \quad (3.5'e)$$

$$\lambda_j(j_{k_o}) \geq 0, \forall j, j_{k_o}; s_{ij_{k_o}}, s_{rj_{k_o}}, d_r^-, d_r^+, d_i^-, d_i^+ \geq 0, \forall i, r, j_{k_o}; \theta_{k_o} \text{ unrestricted} \quad (3.5'f)$$

One can describe the goal achievement variables as adjustments to the level 2 projections that are required in order to permit the sum of projections of the level 1 sub-units to equate to the (adjusted) level 2 projections. If a given pair of goal achievement variables (e.g., d_i^-, d_i^+) are both set to zero (0) in the optimisation of (3.5′), then the goal in question has been met. If, on the other hand, one of those two variables is positive, then the goal in question has not been met, meaning that the sum of the level 1 projections cannot be forced to equate to the level 2 projection.

One can of course contemplate replacing the *difference* between a pair of non-negative goal variables (e.g., $d_i^- - d_i^+$) by a single unsigned variable $d = d^- - d^+$. While this idea works well in the constraint terms, the same is not true for the *sums* of goal variables in the objective function in that the terms such as $d_i^- + d_i^+$ would need to be expressed as $d^- + d^+ = |d|$, where $|d|$ is the modulus or absolute value of d. This resulting non-linearity renders the idea rather undesirable. Even if one accepts this non-linearity there may be a more compelling argument to use the pair of non-negative variables in that in the objective function, one can apply different weights to the under achievement variables d^- than to the over achievement variables d^+. This is not the case if a single unsigned variable is used.

Theorem 3.2. For any given input or output dimension (i or r) at least one member of the pair of deviation variables (e.g., d_i^- *or* d_i^+) must be zero (0), meaning that a goal cannot be both over achieved and underachieved at the same time.

Proof. In the execution of the simplex method the column vectors corresponding to such pairs of deviation variables are linearly dependent, meaning that they cannot be both in the basis at the same time, hence at most one can be positive.

Table 1. Input and output data for sub-units.

		Outputs		Inputs		
Group	Unit	OPER(O1)	OUT(O2)	EQDER(I1)	MAINT(I2)	OCCUP(I3)
Plant 1	1	573	95	110	538	895
	2	560	138	120	290	770
	3	637	151	150	386	886
Plant 2	1	521	102	93	440	771
	2	634	93	102	324	780
	3	610	86	75	378	825
Plant 3	1	620	120	130	350	750
	2	550	81	95	630	770
Plant 4	1	430	105	140	190	810
Plant 5	1	560	110	105	280	770
Plant 6	1	650	170	140	300	7000
	2	550	120	120	275	800
Plant 7	1	320	70	110	230	790
	2	250	60	110	220	790
	3	370	100	140	320	840
Plant 8	1	520	120	100	281	750
	2	430	100	140	302	850
Plant 9	1	475	100	120	179	750
	2	560	150	120	143	800

Table 2. Input and output data for power plants.

Plants	O1	O2	I1	I2	I3
Plant 1	1770	384	380	1214	2551
Plant 2	1765	281	270	1142	2376
Plant 3	1170	201	225	980	1520
Plant 4	430	105	140	190	810
Plant 5	560	110	105	280	770
Plant 6	1200	290	260	575	7800
Plant 7	940	230	360	770	2420
Plant 8	950	220	240	583	1600
Plant 9	1035	250	240	322	1550

This completes the proof.

In the following section, we apply the above-discussed two-stage methodology.

4. Evaluating efficiencies in a hierarchy involving power plants and sub-units

4.1. Analysis factors

As discussed in Cook et al. (1998) and Cook and Green (2005), the following inputs and outputs were used to evaluate the efficiencies of the DMUs and sub-units.

Outputs:

OPER – A function of equivalent full capacity operating hours.

OUT – A function of the number of forced and sudden outages.

Inputs:

EQDER – a function of forced deratings caused by equipment failures.

MAINT – Total maintenance expenditures.

OCCUPT – A function of total occupied hours.

4.2. The results

Table 1 presents the data on the 19 individual sub-units within 9 power plants, for example, Plant #1 contains 3 power units. We point out that the set of power units (19) analysed here is less than the earlier set of 40 units evaluated in the previous articles by Cook et al. (1998) and Cook and Green (2005).

Table 3. Efficiencies and slacks for power plants (Model (3.4)).

Group	Efficiency	S1	S2	S3	S4	S5	Theta
Plant 1	0.919837	0	0	0	420.5771	0	0.919879
Plant 2	0.976185	0	0	0	177.3098	0	0.950020
Plant 3	0.949987	0	0	0	327.009	0	0.950020
Plant 4	0.731708	0	0	13.3522	0	0	0.731709
Plant 5	0.961009	0	0	0	37.68398	0	0.961013
Plant 6	0.901445	0	0	0	121.9699	5340.102	0.901991
Plant 7	0.537672	0	0	0	118.5949	0	0.537683
Plant 8	0.803377	0	0	0	129.0712	0	0.803390
Plant 9	0.935443	0	9.736221	10.43206	0	0	0.935445

Table 4. Sub-unit efficiencies and slacks and plant goal variables (Model (3.5′)).

Group	Unit	Sub-DMU efficiency	S1 (d1)	S2 (d2)	S3 (d3)	S4 (d4)	S5 (d5)
Plant 1	1	0.8643	0.0000	0.0000	0.0000	189.9843	43.4861
	2	0.9933	0.0000	0.0000	0.0000	95.4501	0.0000
	3	0.9618	0.0000	0.0000	8.9930	126.9585	0.0000
	D−		0.0000	0.0000	0.0000	0.0000	0.0000
	D+		0.0000	0.0000	0.0000	15.7773	0.4957
Plant 2	1	0.9280	0.0000	0.0000	0.0000	172.0522	0.0000
	2	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	3	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	D−		0.0000	0.0000	0.0000	0.0000	0.0000
	D+		0.0000	0.0000	6.7964	30.6443	63.2223
Plant 3	1	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.8963	0.0000	0.0000	0.0000	274.7472	0.0000
	D−		0.0000	0.0000	0.0000	0.0000	3.9138
	D+		0.0000	0.0000	1.3898	35.8831	0.0000
Plant 4	1	0.7317	0.0000	0.0000	13.3522	0.0000	0.0000
	D−		0.0000	0.0000	0.0000	0.0000	0.0000
	D+		0.0000	0.0000	0.0000	0.0000	0.0000
Plant 5	1	0.9610	0.0000	0.0000	0.0000	37.6840	0.0000
	D−		0.0000	0.0000	0.0000	0.0000	0.0000
	D+		0.0000	0.0000	0.0000	0.0000	0.0000
Plant 6	1	0.9738	0.0000	0.0000	0.0000	114.3579	5890.3438
	2	0.9617	0.0000	0.0000	17.2099	45.5557	0.0000
	D−		0.0000	0.0000	0.0000	0.0000	0.0000
	D+		0.0000	0.0000	0.0000	0.0000	0.0000
Plant 7	1	0.5733	0.0000	0.0000	0.0000	10.3079	25.0419
	2	0.4591	0.0000	0.0000	0.0000	22.4693	16.9507
	3	0.6349	3.3333	0.0000	8.8889	107.8413	0.0000
	D−		0.0000	0.0000	0.0000	0.0000	0.0000
	D+		3.3333	0.0000	0.0000	0.0000	5.7574
Plant 8	1	0.9754	0.0000	0.0000	0.0000	85.9371	0.3000
	2	0.6701	0.0000	0.0000	2.6106	29.4127	0.0000
	D−		0.0000	0.0000	4.0670	0.0000	0.0000
	D+		0.0000	0.0000	0.0000	21.8128	15.4376
Plant 9	1	0.8693	0.0000	10.9736	9.6943	0.0000	0.0000
	2	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	D−		0.0000	0.0000	0.0000	2.6110	0.0000
	D+		0.0000	1.2374	0.5453	0.0000	2.0249

Table 5. Input and output's *D* value (model 3.5′).

Group		*d*1	*d*2	*d*3	*d*4	*d*5
Plant 1	D−	0.0000	0.0000	0.0000	0.0000	0.0000
	D+	0.0000	0.0000	0.0000	15.7773	0.4957
Plant 2	D−	0.0000	0.0000	0.0000	0.0000	0.0000
	D+	0.0000	0.0000	6.7964	30.6443	63.2223
Plant 3	D−	0.0000	0.0000	0.0000	0.0000	3.9138
	D+	0.0000	0.0000	1.3898	35.8831	0.0000
Plant 4	D−	0.0000	0.0000	0.0000	0.0000	0.0000
	D+	0.0000	0.0000	0.0000	0.0000	0.0000
Plant 5	D−	0.0000	0.0000	0.0000	0.0000	0.0000
	D+	0.0000	0.0000	0.0000	0.0000	0.0000
Plant 6	D−	0.0000	0.0000	0.0000	0.0000	0.0000
	D+	0.0000	0.0000	0.0000	0.0000	0.0000
Plant7	D−	0.0000	0.0000	0.0000	0.0000	0.0000
	D+	3.3333	0.0000	0.0000	0.0000	5.7574
Plant 8	D−	0.0000	0.0000	4.0670	0.0000	0.0000
	D+	0.0000	0.0000	0.0000	21.8128	15.4376
Plant 9	D−	0.0000	0.0000	0.0000	2.6110	0.0000
	D+	0.0000	1.2374	0.5453	0.0000	2.0249

Table 2 displays the aggregates of the data on sub-units within each of the nine plants. As an example, the value for output #1 for Plant 1 is 1770, which is the sum of the three corresponding values (573, 560 and 637) for that variable held by sub-units 1,2 and 3, respectively within Plant 1.

Microsoft Solver was used to generate the efficiency scores in stage 1 for the 9 power plants (using model (3.4)), as displayed in Table 3, along with the associated slacks. The value ofε used herein was 0.00001. The column labelled "Theta" records the value of$\hat{\theta}$, the CCR DEA weak efficiency portion of the overall score ("Efficiency"). Plant 8, for example, has a$\hat{\theta}$-value of 0.803390, but an overall efficiency of 0.803377 when nonzero slacks of 129.0712 for input #2 are taken into account. This means that a reduction of 19.6610% in the three inputs projects DMU 8 to a *weak* efficiency point (not on the frontier) that requires further adjustments in terms of the slacks in order to reach the actual frontier. Overall, none of the plants are strongly efficient (lie on the frontier).

Table 4 presents the DEA efficiency scores $\hat{\theta}$(arising from model (3.5′)) and associated slacks for the 19 sub-units within the 9 plants. In addition, the goal achievement variables $D-, D+$ are displayed. We note that four of the sub-units are strongly DEA efficient (#2 and #3 in plant 2, #1 in plant 3 and #2 in plant 9). Note as well that 3 of the plants (4–6) have weakly inefficient sub-units (have non-zero slacks), but have zero goal variables, meaning that model (3.5) is feasible for these plants under Principal 2 are met, hence the goals are achieved (Table 5).

5. Conclusions

This paper has re-examined the earlier research by Chai et al. (1998) and Cook and Green (2005) involving the performance of electric power plants. We develop a model that looks at two-level hierarchical structures for those particular settings where the performance data of the second level DMU (the power plant) is explicitly defined as being the aggregate of the corresponding data on inputs and outputs held by the level 1 sub-units within the plant. We see this development as an important first step toward more general situations.

One direction is aimed at situations where level 2 DMUs (in the current setting, power plants) have their own inputs and outputs, separate from the factors at the sub-unit level. In Cook and Green (2005) one such example of this was raised, namely CO_2 emissions that are explicitly connected to the plant and not the sub-units within the plant. As discussed earlier, one solution to this is to split the factor by some means and allocate the resulting amounts to the various sub-units.

Another direction involves the presence of data at the sub-unit level that do not readily lend themselves to aggregation. An example of this might be the number of years of experience of sub-unit management, where aggregation would not apply. Similarly, age of power units would not be a variable allowing for aggregation. One possible approach is to recognise that in model (3.3) one can replace, for example, the *aggregate* input $X_{ik_o} = \sum_{j_{k_o} \in J_{k_o}} x_{ij_{k_o}}$(or outputs) by the *average* of the sub-unit values $\bar{x}_{ik_o} = \sum_{j_{k_o} \in J_{k_o}} x_{ij_{k_o}} / |K_k|$, where $|K_k|$ denotes the *cardinality* of the set of sub-units K_k in group (e.g., power plant) k. Thus, one can replace, for example, the aggregate age of sub-units by the average age of those sub-units or replace the aggregate of rank order data or Likert scales by the average of those scales. At the same time, it may be necessary to correct for the fact that the averages of say Likert scale data values may not lie on such a scale (see, e.g., Du, Chen, Cook, Hu, and Zhu 2017).

Acknowledgement

The authors wish to thank two anonymous referees and an Associate Editor.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Wade Cook and Zhepeng Li were supported under grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) (A8966 and RGPIN-2017-05667). Joe Zhu acknowledges support from the National Natural Science Funds of China (No. 71828101).

ORCID

Joe Zhu http://orcid.org/0000-0001-6103-080X

References

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, *30*(9), 1078–1092. doi:10.1287/mnsc.30.9.1078

Castelli, L., Pesenti, R., & Ukovich, W. (2004). DEA-Like models for efficiency evaluations of hierarchically structured units. *European Journal of Operational Research*, *154*(2), 465–476. doi:10.1016/S0377-2217(03)00182-6

Chai, D., Cook, W. D., Green, R. H., & Doyle, J. (1998). Hierarchies and groups in DEA. *Journal of Productivity Analysis*, *10*(2), 177–198. doi:10.1023/A:1018625424184

Charnes, A., Cooper, W. W., Golany, B., Seiford, L. M., & Stutz, J. (1985). Foundations of data envelopment analysis and Pareto-Koopmans empirical production functions. *Journal of Econometrics*, *30*(1–2), 91–107. doi:10.1016/0304-4076(85)90133-2

Charnes, A., Cooper, W. W., & Rhodes, E. L. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, *2*(6), 429–444. doi:10.1016/0377-2217(78)90138-8

Cook, W. D., & Green, R. H. (2005). Evaluating power plant efficiency: A hierarchical model. *Computers & Operations Research*, *32*(4), 813–823. doi:10.1016/j.cor.2003.08.019

Cook, W. D., & Seiford, L. M. (2009). Data envelopment analysis (DEA)-thirty years on. *European Journal of Operational Research*, *192*(1), 1–17. doi:10.1016/j.ejor.2008.01.032

Du, J., Chen, Y., Cook, W. D., Hu, H., & Zhu, J. (2017). Bounded and discrete data and Likert scales in data envelopment analysis: Application to regional energy efficiency in China. *Annals of Operations Research*, *255*, 347–366. doi:10.1007/s10479-015-1827-3

Kao, C. (2009). Efficiency measurement of parallel production systems. *European Journal of Operational Research*, *196* (3), 1107–1112. doi:10.1016/j.ejor.2008.04.020

Kao, C. (2015). Efficiency measurement for hierarchical network systems. *Omega*, *51*, 121–127. doi:10.1016/j.omega.2014.09.008

Li, W., Cook, W. D., Li, Z., & Liang, L. (2017). Evaluation of ecological systems and the recycling of undesirable outputs: An efficiency study of regions in China. *Socio-Economic Planning Sciences*, *60*, 77–86. doi:10.1016/j.seps.2017.03.002

Liang, L., Cook, W. D., & Zhu, J. (2010). Measuring performance of two-stage network structures by DEA: A review and future perspectives. *Omega*, *38*(6), 423–430. doi:10.1016/j.omega.2009.12.001

Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research*, *130*(3), 498–509. doi:10.1016/S0377-2217(99)00407-5